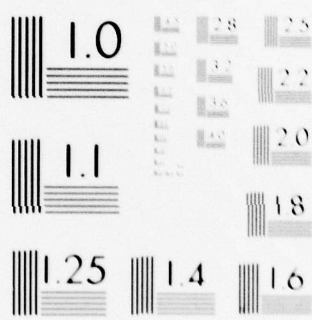
AD-A075 535 MINNESOTA UNIV MINNEAPOLIS DEPT OF PSYCHOLOGY F/G 5/10
EFFECT OF POINT-IN-TIME IN INSTRUCTION ON THE MEASUREMENT OF AC--ETC(U)
AUG 79 G G KINGSBURY , D J WEISS N00014-76-C-0627
UNCLASSIFIED RR-79-4 NL

1 OF 1
AD
A075535

END
DATE
FILMED
3-80
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ADA075535

EFFECT OF POINT-IN-TIME IN INSTRUCTION ON THE MEASUREMENT OF ACHIEVEMENT

G. Gage Kingsbury
and
David J. Weiss

RESEARCH REPORT 79-4
AUGUST 1979

PSYCHOMETRIC METHODS PROGRAM
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN 55455

This research was supported by funds from the Navy Personnel Research and Development Center, Army Research Institute, Air Force Human Resources Laboratory, Defense Advanced Research Projects Agency, and the Office of Naval Research, and monitored by the Office of Naval Research.

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

79 22 22 120

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Research Report 79-4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Effect of Point-in-Time in Instruction on the Measurement of Achievement		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) G. Gage Kingsbury and David J. Weiss		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0627
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology University of Minnesota Minneapolis, MN 55455		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS P.E.: 61153N PROJ.:RR042-04 T.A.: RR042-04-01 W.U.: NR150-389
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research Arlington, VA 22217		12. REPORT DATE August 1979
		13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This research was supported by funds from the Navy Personnel Research and Development Center, Army Research Institute, Air Force Human Resources Laboratory, Defense Advanced Research Projects Agency, and the Office of Naval Research, and monitored by the Office of Naval Research.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

achievement testing	pretest-test measurement	change scores
latent trait test theory	test-posttest measurement	adaptive testing
item response theory	measurement of change	tailored testing
item characteristic curve theory	measurement of growth	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Item characteristic curve (ICC) theory has potential for solving some of the problems inherent in the pretest-test and test-posttest paradigms for measuring change in achievement levels. However, if achievement tests given at different points in the course of instruction tap different achievement dimensions, the use of ICC approaches and/or change scores from these tests is not desirable. This problem is investigated in two studies designed to determine whether or not achievement tests administered at different times

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

during a sequence of instruction actually measure the same achievement dimensions.

To investigate possible changes in dimensionality between different points in instruction, aspects of the dimensionality of achievement test data were examined prior to instruction, at the peak of instruction, and up to a month following the peak of instruction. Data used were conventional and adaptive achievement test data administered to students in a general biology course at the University of Minnesota.

Results raised questions about the utility of the pretest-test paradigm for measuring change in achievement levels, since a comparison of ICC parameter estimates indicated that a change in the dimensionality of achievement had occurred within the short (4-week) period of instruction. This change was also observed using a factor analytic comparison.

Use of the test-posttest paradigm to measure retention was supported, since a regression comparison of students' achievement level estimates did not indicate any significant change in the achievement metric up to 1 month after the peak of instruction. The significance of this result for the use of adaptive testing technology in measuring achievement is described.

Implications of these studies and the use of ICC theory in the measurement of achievement, as well as some potential limitations in terms of generalizability of these results, are discussed.

CONTENTS

Acknowledgments

Adaptive and conventional test data utilized in this report were obtained from volunteer students in Biology 1-011, General Biology, at the University of Minnesota during fall quarters 1976 and 1977 and winter quarter 1977; appreciation is extended to these students for their participation in this research. The cooperation of Kathy Swart and Norman Kerr of the General Biology staff in providing access to the students is also deeply appreciated.

Technical Editor: Barbara Leslie Camm

Effect of Point-in-Time in Instruction on the Measurement of Achievement

The measurement of achievement is a considerably more complex problem than the measurement of ability. Whereas ability levels develop over long periods of time and remain relatively stable with exposure to different environments, achievement levels characteristically result from exposure to specific instructional or training environments. Although these instructional environments may span relatively long periods of time, such as the 2- to 4-year periods required for training in many skilled fields, instructional decisions are more typically made on the basis of instructional periods of several weeks or a few months. In the extreme case, as in computer-assisted instruction, the instructional environment designed to modify individuals' achievement levels may be as short as a few minutes. Thus, achievement is a dynamic variable which may change over very short time intervals.

The Pretest-Test Paradigm

Because achievement levels should be sensitive to instruction, it may be desirable to measure achievement (1) prior to instruction, (2) at the end (or peak) of instruction, and (3) some time after the completion of instruction. The measurement of achievement prior to instruction is accomplished by means of pretests, designed to determine a student's level of achievement before exposure to the instructional environment. To determine whether instruction has had an effect, a pretest-test paradigm may be used, in which group or individual gain scores are computed to demonstrate the effects of the instruction.

The pretest-test paradigm for measuring achievement has at least two major problems. First, both individual and group change scores have been shown to be highly unreliable (Cronbach & Furby, 1970) unless the pretest and peak-of-instruction test measurements are each extremely precise. Second, it is necessary to administer two tests covering the same material to students. If the same test is administered prior to instruction and at the peak of instruction, students may not be motivated to respond optimally to both tests; thus, precision of measurement on the tests may be lowered. At the same time, students may remember test items from the pretest, find the answers to those items in the course materials, and then perform better on the second test than they would have performed had they not seen the test items prior to instruction.

On the other hand, if different tests are used prior to instruction and at the peak of instruction, the serious problem of developing parallel tests arises. This is compounded by the necessity to obtain very precise measurements, which may result in pretests and peak-of-instruction tests measuring different aspects of the achievement variable in the quest for highly precise measurements.

A potential solution to these problems in the pretest-test paradigm lies in the application of techniques of item characteristic curve (ICC) theory

(Lord & Novick, 1968) to the measurement of achievement. Achievement test items calibrated using ICC theory (e.g., Bejar, Weiss, & Kingsbury, 1977) will all be on the same metric. Thus, selection of ICC-calibrated items from the same item pool to constitute pretests and peak-of-instruction tests will, in theory, eliminate the need for the construction of parallel tests. In addition, placing all the items on the same metric by using ICC item parameters will eliminate the need to repeat the same items at the two testings, since (again in theory) any subset of items from the precalibrated pool will measure the same variable as any other subset of items. Thus, items for pretests and for peak-of-instruction tests can be selected from the ICC-calibrated pool on the basis of content considerations resulting in effectively parallel measurements.

ICC theory can also be applied to the pretest-test paradigm of achievement measurement through the use of adaptive testing to increase the precision of the achievement measurements, thus possibly permitting the use of individual or group gain scores. Research by Bejar and Weiss (Bejar & Weiss, 1978; Bejar, Weiss, & Gialluca, 1977) in an achievement testing context shows that ICC-based adaptive tests produce measurements which are considerably more precise than those of conventional achievement tests, supporting similar findings in the ability-testing literature (e.g., McBride & Weiss, 1976; Vale, 1975; Vale & Weiss, 1975).

Before ICC theory can be applied in the pretest-test paradigm of achievement measurement, however, it must be demonstrated that data obtained in this paradigm meet the assumptions of the theory. Specifically, since most ICC-based techniques require unidimensionality, it must be shown that both pretests and peak-of-instruction tests are essentially unidimensional or that, in general, the dimensionality of the two tests is the same. In addition, it must be demonstrated that the latent space in the two tests does not change. That is, even though both the pretest and the peak-of-instruction test are unidimensional, it is possible that they are measuring achievement on different dimensions. If this is the case, item parameters estimated at one point in instruction would not be usable at the other point in instruction.

The Test-Posttest Paradigm

Just as measured achievement is expected to change in level from pretest to peak of instruction, it is also expected that it should remain stable for some time after instruction. Thus, it is appropriate to investigate whether measured achievement levels deteriorate over short or long periods of time in order to determine the permanency of the instructional effect demonstrated by the pretest-test data. Such a demonstration would require the test-posttest paradigm in which the peak-of-instruction test is followed at some point in time by the administration of a posttest.

Because the test-posttest paradigm for measuring constancy of achievement may be implemented with either the parallel tests approach or the repeated tests approach, it has exactly the same problems as the pretest-test paradigm. Similarly, the use of ICC theory and adaptive testing may be brought to bear on these problems if the peak-of-instruction and posttest data meet the requirements of these approaches. Thus, similar kinds of data must be generated to investigate the use of these approaches in an achievement context.

The use of adaptive testing in the measurement of achievement--whether at pretest, peak of instruction, or posttest--raises an additional problem which

requires investigation. To realize the potential gains in the measurement of achievement in increased precision (Bejar, Weiss, & Gialluca, 1977), higher validity (Bejar & Weiss, 1978), and shorter testing times (Brown & Weiss, 1977), adaptive tests should be administered by computer. In achievement environments with large numbers of students, there may not be sufficient numbers of adaptive testing terminals so that each student can be tested at the peak of his or her instruction. Thus, it may be necessary for students to have their achievement measured at some point beyond the peak of instruction. A similar situation exists in self-paced instructional environments, where students may not take achievement tests exactly at the peak of instruction due to procrastination or influences beyond their control (e.g., unavailability of equipment). In both cases, it is an important question whether achievement measured after the peak of instruction is measured on the same dimension as achievement measured at the peak of instruction.

Objectives

The studies reported below were designed to investigate several questions relevant to the implementation of ICC theory in pretest-test and test-posttest paradigms for measuring achievement. The data also have some bearing on the practical questions involved in the use of adaptive testing in measuring achievement within the realistic constraints of instructional environments.

STUDY 1: RELATIONSHIP OF TEST CHARACTERISTICS PRIOR TO INSTRUCTION AND AT PEAK OF INSTRUCTION

This study was designed to investigate two questions concerning test characteristics of a test used to measure achievement at the peak of instruction when it was applied to a population of testees measured prior to instruction:

1. Are ICC item parameters estimated from data obtained prior to instruction quantitatively equivalent to parameters estimated at the peak of instruction? This question is concerned with whether the ICC metric maintains its interval properties during the course of instruction.

2. Do tests used to measure achievement prior to instruction (i.e., pretests) measure attributes from the same latent space as tests used to measure achievement at the peak of instruction? This question is concerned with whether the responses of the two populations (pretest versus test) can be described by a common latent space.

If the responses to both of these questions are affirmative, the results may be taken as support for the pretest-test paradigm. These results would also have implications for the power of the unidimensional ICC model for measuring achievement during the course of instruction. If major differences are found in the characteristics of the tests used to measure achievement prior to and at the peak of instruction, the foundation of the pretest-test paradigm for measuring achievement would be weakened in many applications and the use of the ICC model to measure growth in achievement would be limited.

Method

Test Data

Prior to instruction. The testing sessions that provided the prior-to-instruction data were classroom examinations administered on the first day of class during the fall academic quarter of 1977 to all students attending class for Biology 1-011, General Biology, at the University of Minnesota. (For a description of the course and testing procedures, see Bejar, Weiss, & Kingsbury, 1977.) Data were obtained from 1,294 students. The test administered at this time consisted of 40 multiple-choice items sampled from all of the 7 content areas covered in the course; these items were taken from a larger pool of items developed for this course.

Peak of instruction. The peak-of-instruction test data were obtained from two sources so that two different types of questions could be answered. The first question concerned whether item parameter estimates obtained from prior-to-instruction testing were similar to estimates obtained for the same items at peak-of-instruction testing. Peak-of-instruction parameter estimates were used that were obtained from test data supplied by individuals enrolled in the same course during five earlier academic quarters, since it would be inappropriate to administer the same items twice to the same individuals. These calibration samples averaged between 700 and 1,000 students. (For the exact number of subjects responding to each of the items for calibration, see Kingsbury & Weiss, 1979, p. 26.)

The second question concerned whether the factor structure underlying students' responses changed as a function of instruction. To answer this question, student reponse data were used which were collected on a 55-item midquarter examination administered 4 weeks after the pretest, as one of the course requirements, to approximately the same group of students who took the pretest. Approximately 1,200 students completed the 55-item midquarter examination. Each student was required to omit 5 items in the examination; consequently, data for each item were based on about 1,000 students.

Item Parameter Analysis

Item parameterization. Estimation of ICC item parameters for the peak-of-instruction data is described in detail in Bejar, Weiss, and Kingsbury (1977). In brief, a computer program developed by Urry (1976) was used to fit a three-parameter logistic ogive for each item administered to the testees. Items were rejected by the parameter estimation program if they failed to reach certain minimal standards with respect to their discrimination value (a) and lower asymptote (c). Thus, values for the index of discriminatory power (a), (b), and probability of attaining a correct answer with no knowledge of the subject (c) were obtained for each item that surpassed the minimum standards. Specifically, an item was rejected if during the first stage of the parameter estimation process the value of its a parameter estimate was less than .80 or the value of its c parameter estimate was greater than .30.

This procedure was applied separately to both the prior-to-instruction data and the peak-of-instruction data. The results, for each of the 40 items administered at both points in instruction, were two comparable estimates of the ICC parameter values, varying only because of sample fluctuation and the difference in the instructional level of the two groups.

Comparison of item parameter estimates. If the latent space is constant, the parameter estimates from the prior-to-instruction and the peak-of-instruction groups should differ no more than the estimates obtained from two samples at the same level of instruction. To provide a basis of comparison for the prior-to-instruction versus peak-of-instruction correlations, ICC item parameter estimates were computed from two groups of students in the same course during two earlier quarters who answered a comparable group of items at the peak of instruction (the peak/peak group). The peak/peak data were partially reported before by Bejar, Weiss, and Kingsbury (1977), who reported the peak/peak correlations for the a and b parameter estimates for 18 items obtained from responses of approximately 900 testees in each sample. Appendix Table A shows item numbers and parameter estimates of the items used to investigate sampling variation in parameter estimates obtained from the two groups at the peak of instruction. Also in this table are the times of administration of the items to the students.

To compare ICC parameter estimates obtained prior to instruction with those obtained at the peak of instruction, Pearson product-moment correlations were computed between item parameter estimates obtained at the two time periods (prior/peak correlations) separately for the a, b, and c parameters. The three correlations obtained were also computed in the two samples which were at the peak of instruction (peak/peak correlations). For each of the three parameters, the prior/peak and peak/peak correlations should differ only to the extent that the individuals differed in their test performance when the group was tested prior to instruction compared with their performance when tested at the peak of instruction, as reflected in the ICC item parameters.

To determine whether the prior/peak correlations differed significantly from the peak/peak correlations, Fisher's z'-transformation was applied to the prior/peak correlations and a confidence interval was constructed around each correlation (Neter & Wasserman, 1974). If these intervals included the observed peak/peak correlations, the hypothesis that the obtained correlations might come from the same population could not be rejected. If a confidence interval around the prior/peak correlation did not include the value of the observed peak/peak correlation, it could be concluded that the differences between the observed correlations were probably not due to sampling fluctuation. If the peak/peak correlation fell above the upper limit of the prior/peak confidence interval, this would imply that the ICC parameters were not invariant between the prior-to-instruction sample and the peak-of-instruction sample. This variability of parameter values would indicate that the two samples reflected different populations and that the ICC parameters estimated in the peak-of-instruction population were not sufficient to describe the responses of individuals in the prior-to-instruction population.

Factor Structure

Of the items administered in the pretest, 21 were sampled from the content areas taught in the first portion of the course; these content areas were then tested on the first midquarter examination, which was administered later in the course. The items were used to investigate the factor structure prior to instruction.

Twenty-one items tapping the same content areas were chosen arbitrarily from the first midquarter examination, which was administered to the same individuals 4 weeks after the pretest. The students' responses to these items

were used to examine the factor structure underlying performance at the peak of instruction. Items administered at the pretest and at the first midquarter were sampled from the same content areas, but different items were used at the two points in time.

For each of these groups of items, the same procedure was followed to obtain the final factor structure. First, all student responses were scored "0" if incorrect or "1" if correct. Second, these recoded responses were used to obtain tetrachoric correlations among the items through the TETRACHORIC subroutine in the Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1970). The two resultant correlation matrices were then factor analyzed using the FACTOR subroutine from the same statistical package. The final factor solutions were obtained using a principal axis solution; the initial communality estimates were the squared multiple correlations of each variable with all the other variables. The factor solutions, which were arbitrarily limited to five factors, were iterated until the differences in successive communality estimates were negligible. This procedure provided the final solutions.

The two final factor solutions were then compared for similarities and differences in terms of the number of salient factors, the strength of each factor, and the amount of variance in the item intercorrelations accounted for by the factor solutions. To the extent that observed differences between the two solutions were minor, it could be inferred that the underlying factors contributing to testee responses were the same prior to instruction and at the peak of instruction. To the extent that major discrepancies were observed between the solutions, it could be inferred that differences existed in the structure of achievement at the two points in instruction.

Results

Item Parameters

Parameter estimates for each of the 40 items administered in the prior-to-instruction achievement measure are shown in Table 1, along with parameter estimates for the same items obtained from groups of testees at the peak of instruction. From this table it can be seen that 14 of the items failed to meet the minimal standards of the estimation procedure when administered prior to instruction and 5 items failed when administered at the peak of instruction. Four items were rejected in both instances. After all of the rejected items were removed from consideration, 25 items remained for the correlational analysis.

The bivariate plots of a, b, and c parameter estimates from the two calibrations are shown in Figures 1, 2, and 3, respectively. It can be seen from these figures that the relationships between the sample estimates of the parameter values were weak, at best. Figure 1 shows a correlation of -.12 for the a parameter, indicating a slight tendency for high values of a at peak of instruction to be associated with low values prior to instruction.

The correlation of r=.64 for the b parameter data show a tendency for high values of b prior to instruction to be associated with high values at the peak of instruction. However, almost all the data points in Figure 2 are below the main diagonal, indicating a tendency for items to be more difficult prior to

Table 1
Parameter Estimates of Items Calibrated
Prior to Instruction and at the Peak of Instruction

Item	Prior to Instruction			Peak of Instruction		
Number	*a*	*b*	*c*	*a*	*b*	*c*
3035	1.21	2.53	.28	.90	.68	.28
3241	1.89	2.87	.48	.91	2.09	.17
3816	1.72	-.26	.51	--	--	--
4013	1.02	.79	.38	1.76	-1.88	.16
3809	1.12	.70	.35	1.27	-.61	.53
4010	1.25	.00	.43	.88	-1.82	.23
3817	--	--	--	--	--	--
3803	--	--	--	--	--	--
3210	--	--	--	1.04	-1.22	.40
3837	1.10	1.13	.29	1.09	-1.59	.25
3235	1.17	1.38	.49	1.15	-1.40	.28
3808	--	--	--	.99	-1.00	.30
4033	--	--	--	.90	2.23	.38
3812	.85	1.92	.33	.82	-.63	.13
3424	--	--	--	--	--	--
3821	--	--	--	.90	-.92	.43
3244	1.00	2.52	.47	1.35	-.44	.23
3013	.91	1.18	.38	1.00	-.97	.39
3065	1.55	.03	.46	1.17	-1.66	.39
3909	--	--	--	1.34	.77	.38
3922	.76	2.33	.28	.64	-.26	.30
3415	--	--	--	.85	-.96	.41
3428	--	--	--	.90	-1.56	.40
3067	.98	1.15	.33	1.07	-.76	.21
3272	.83	1.45	.40	1.06	-.81	.37
3908	--	--	--	1.15	.07	.31
3435	1.85	1.43	.39	.83	-.61	.42
4005	--	--	--	--	--	--
3426	1.03	2.68	.44	.68	.07	.22
3031	.75	2.54	.24	1.47	-.33	.39
4006	1.01	2.53	.47	.84	-.59	.16
3069	.85	1.11	.45	.88	-.01	.48
3211	--	--	--	.88	.01	.13
3905	--	--	--	.98	.35	.20
4015	.76	2.26	.31	2.03	-1.62	.12
3403	.93	1.38	.33	.99	.18	.19
3000	3.06	2.46	.21	1.24	.52	.36
3445	.73	2.50	.39	1.19	.44	.34
3218	2.44	2.18	.30	.82	.58	.12
4001	1.44	2.49	.29	1.47	-1.14	.13

Note. Missing values indicate that item was rejected by the item calibration procedure.

instruction than at the peak of instruction. The data in Table 1 indicate that prior to instruction only one of the item *b* values was negative (an easy item), but at peak of instruction more than half the items had negative *b* values.

Figure 1
ICC Discrimination (a) Parameter Values Estimated
Prior to Instruction and at Peak of Instruction (r=-.12)

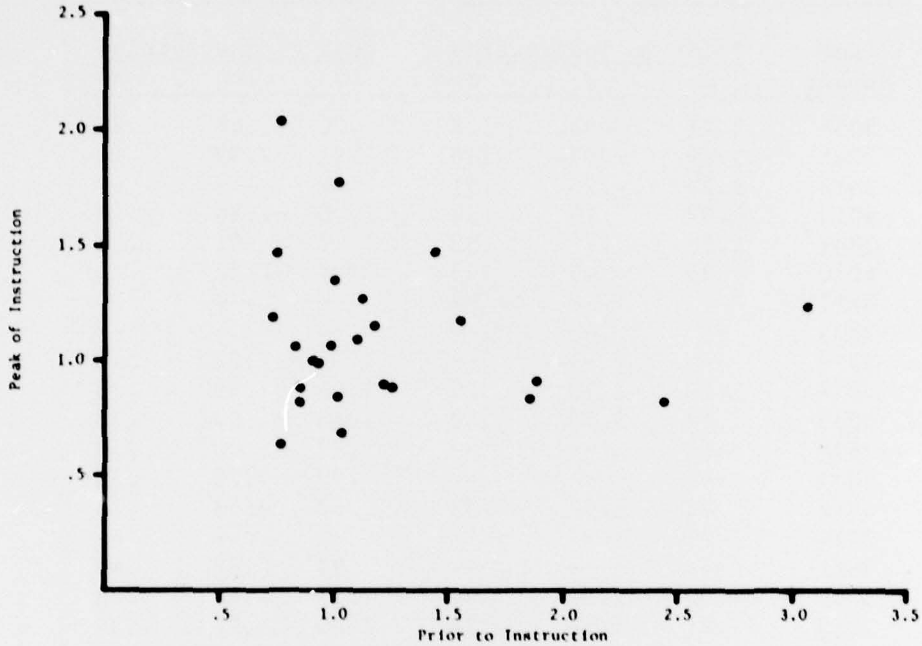


Figure 2
ICC Difficulty (b) Parameter Values Estimated
Prior to Instruction and at Peak of Instruction (r=.64)

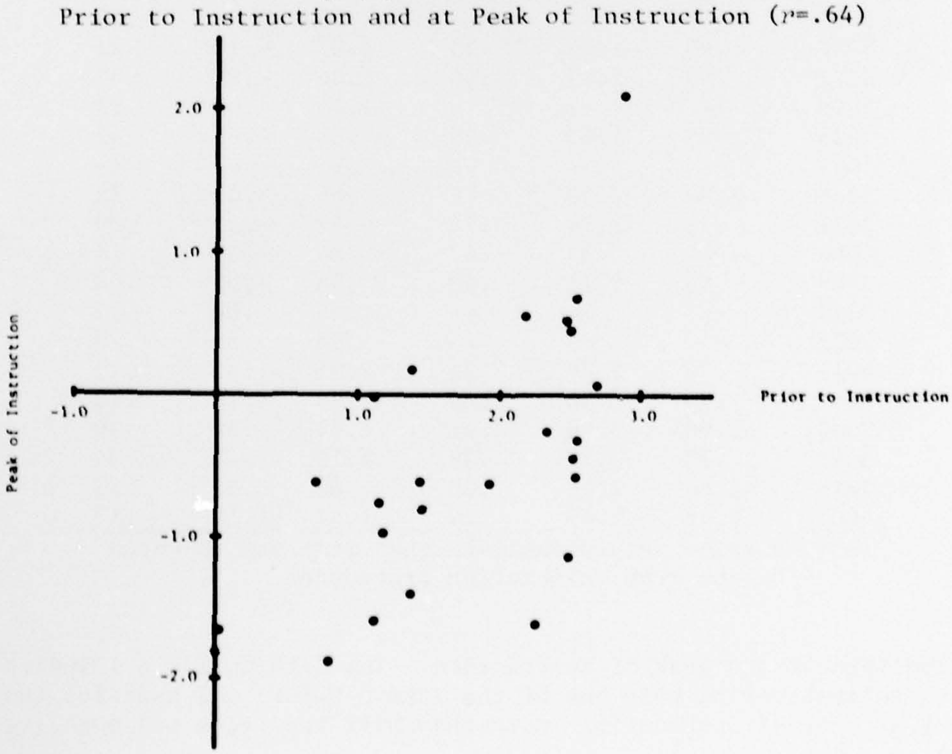
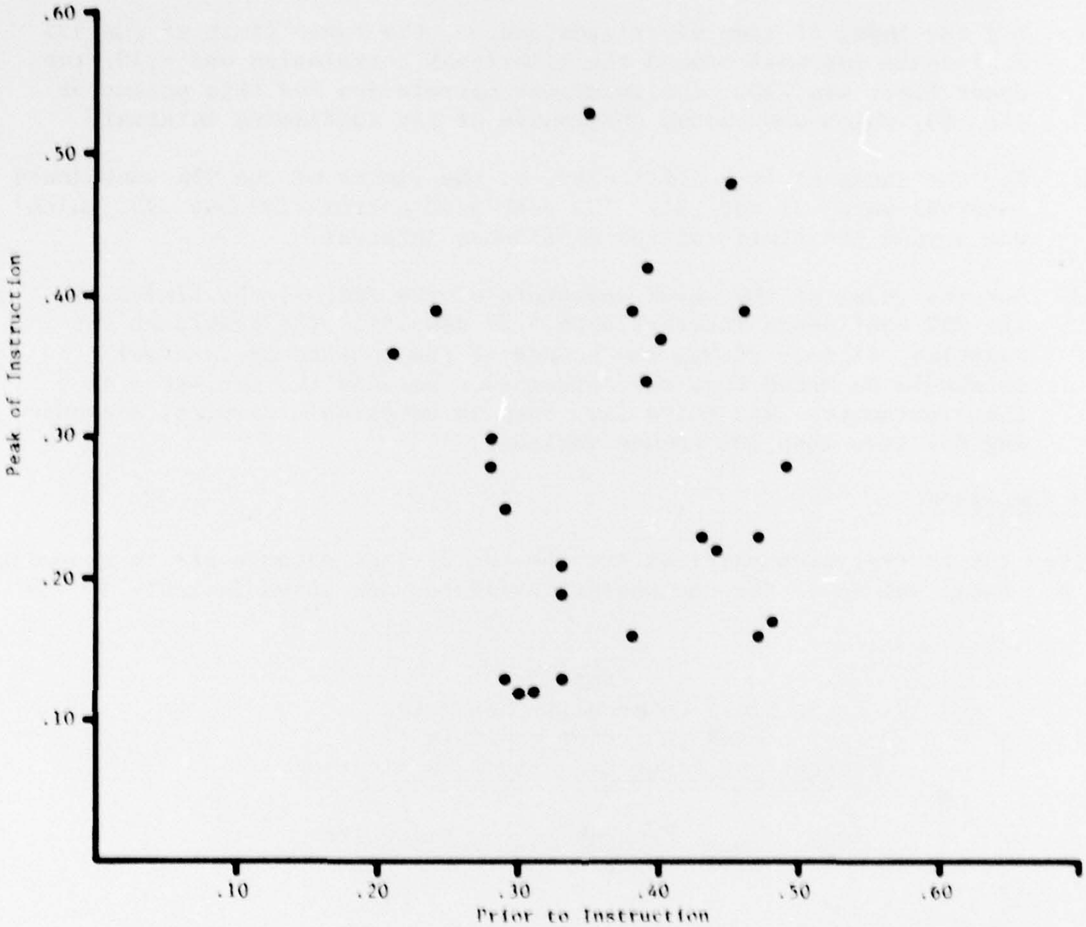


Figure 3
ICC Lower Asymptote (c) Parameter Values Estimated Prior to Instruction
And at Peak of Instruction (r=.04)



The data in Figure 3 indicate essentially no relationship (r=.04) between the c parameters estimated at the two points in time. As expected, however, there was a general tendency for guessing parameter values to be higher prior to instruction than at the peak of instruction.

Table 2
Correlations Between Item Parameter Estimates for
Prior/Peak and Peak/Peak Data

Data	a	b	c
Prior/Peak	-.12	.64	.04
Peak/Peak	.63	.96	.41

The prior/peak correlations (based on item parameter data in Table 1) are shown in Table 2, along with the peak/peak correlations (based on data in Appendix Table A) obtained from separate samples on other items drawn from the same

testing pool. Using the z'-transformation, 95% confidence intervals were computed for each of the prior/peak correlations with the following results:

1. For the index of item discrimination, a, the lower limit of the 95% confidence interval around the prior/peak correlation was -.50; the upper limit was .30. The peak/peak correlation for this parameter was .63, which was beyond the bounds of the confidence interval.

2. For the index of item difficulty, b, the limits of the 95% confidence interval were .31 and .81. The peak/peak correlation was .96, which was beyond the limits of the confidence interval.

3. For the index of the lower asymptote of the ICC, c, the limits of the 95% confidence interval were -.37 and .55. The peak/peak correlation, .41, fell within the bounds of the confidence interval. It should be noted that the correlation between the estimates of the c-parameter was quite low, even in comparable samples, accounting for less than 20% common variance.

Factor Structure

Item intercorrelation matrices for the two 21-item subsets are in Appendix Table B. Final values of the communality estimates are shown in Table 3. The

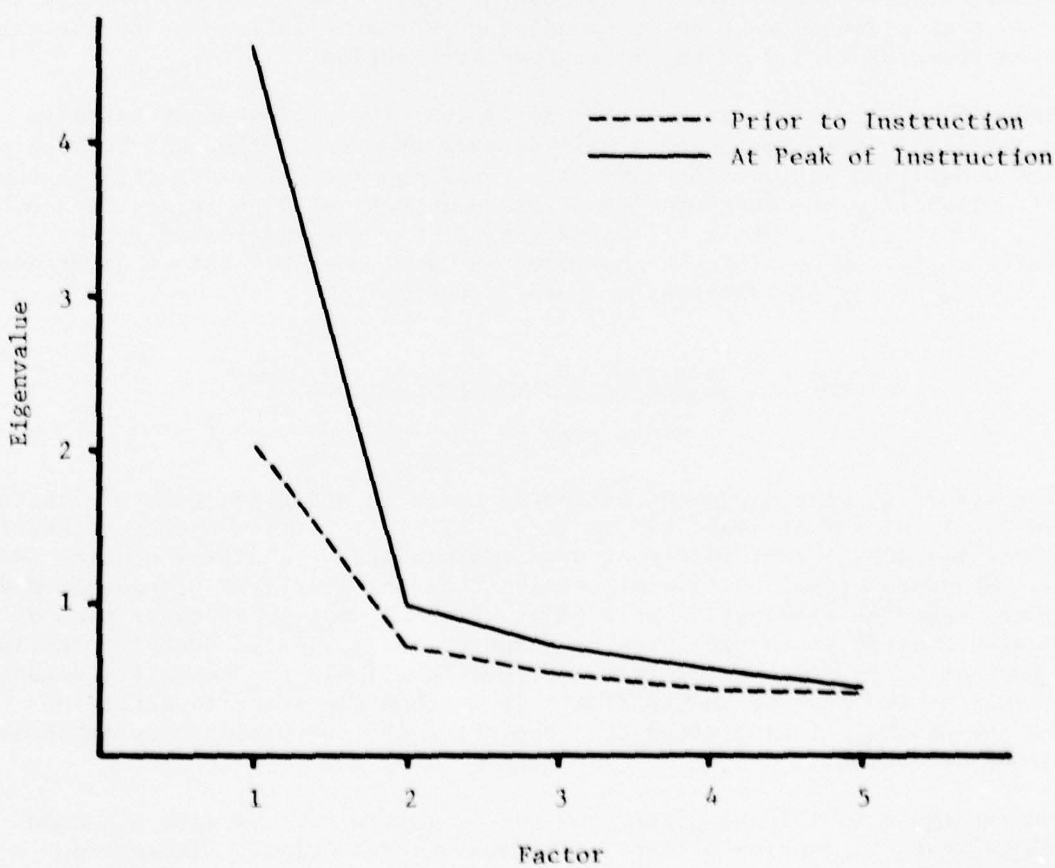
Table 3
Final Communality Estimates
Used in Factor Analysis of
Pretest and First Midquarter Examination

Item	Pretest	First Midquarter
1	.351	.929
2	.024	.166
3	.191	.233
4	.476	.140
5	.127	.960
6	.052	.411
7	.198	.134
8	.378	.218
9	.068	.279
10	.079	.495
11	.242	.345
12	.093	.250
13	.709	.198
14	.031	.451
15	.225	.346
16	.379	.487
17	.110	.192
18	.265	.188
19	.092	.553
20	.034	.365
21	.171	.114

mean communality estimate for the pretest items prior to instruction was .20; the mean communality estimate for the first midquarter items at the peak of instruction was .34. About one-third of the variance for the average test item at the peak of instruction was accounted for by the common factors, compared to one-fifth of the variance for the average item prior to instruction. Consequently, there was more unique variance in the items administered prior to instruction than in the items administered at the peak of instruction.

Eigenvalues for the final factor solutions are presented graphically in Figure 4 (factor loadings are in Appendix Table C). It can be seen that subsequent to the first factor, the eigenvalues for each factor were quite similar. For the first factor, though, the eigenvalue obtained from the items administered at the peak of instruction (4.67) was more than twice the value obtained from those administered prior to instruction (2.04). Thus, although both sets of items had a dominant first factor, in the peak-of-training data a stronger first factor was evident; in the prior-to-training data a weak single factor accounted for achievement.

Figure 4
Eigenvalues of Factors Obtained Prior to Instruction (Pretest) and at Peak of Instruction (First Midquarter Examination)



Conclusions

The correlational analysis indicated that the ICC parameters obtained from two samples differing in instruction were more discrepant than was to be expected from observing correlations among parameters obtained from samples at the same point in instruction. Further, for two of the ICC parameters--discrimination (a) and difficulty (b)--the prior/peak correlations were found to be significantly smaller than the peak/peak correlations. These findings imply that the latent space underlying testee responses changes enough between the beginning and the peak of instruction so that the test responses cannot be described by a single latent continuum; items change, not only with respect to their ability to differentiate testees at different levels of the trait continuum (discrimination), but also with respect to their relative difficulties. The findings imply that in order to describe test item responses obtained both prior to and at the peak of instruction with a single latent trait model, the unidimensional model that was considered adequate to describe performance at the peak of instruction would need to be expanded to a multidimensional model developed from both sets of data.

This conclusion is supported by the comparative factor analyses. These analyses showed a major difference in the strength of the first factor underlying testee responses in the few weeks of instruction between the pretest and the peak of instruction. Students' test item responses were not, to as great an extent, related to the first factor during the pretest. Again, the implication is that students were not responding to the same influences to the same degree on the pretest and on the midquarter examination.

These findings, taken as a whole, imply that the pretest-test paradigm may be invalid in some instances simply because the tests might not be tapping the same underlying achievement variable. This may account, in part, for the lack of reliability of change scores reported in many studies (e.g., Cronbach & Furby, 1970; Harris, 1963). It is suggested that the underlying factor structures in testee responses be explored whenever possible before importance is attributed to any pretest-test measure of change.

STUDY 2: STABILITY OF ACHIEVEMENT ESTIMATES AFTER THE PEAK OF INSTRUCTION

The stability of achievement estimates measured after the peak of instruction is important for at least two reasons. First, it is frequently necessary to measure the achievement levels of some individuals at a different time than others. Students occasionally miss examinations for a variety of reasons and, therefore, take the examination at a point which may not be at their peak of instruction. Where tests are given by computers (e.g., as in adaptive testing), there may not be sufficient terminal equipment available to test all students immediately at the peak of instruction. It is thus important to determine whether the passage of time after the completion of instruction affects achievement level estimates.

Measuring an individual after the peak of instruction is also a common problem in research studies attempting to measure retention. Similar to the pretest-test paradigm, the test-posttest paradigm used to measure reten-

tion assumes that the same achievement variable is being measured and that the passage of time does not change the nature of achievement. Thus, it is again relevant to determine whether achievement level estimates obtained after the peak of instruction are systematically related to those obtained at the peak of instruction.

If it is hypothesized that performance of individuals tested some time after the peak of instruction is a function of the same latent space that influences performance at the peak of instruction, several outcomes would be expected when individuals' achievement levels estimated from tests given at different times are compared. If the unidimensional latent space remains static with the passage of time, achievement estimates for individuals measured at different points in time after the peak of instruction should differ only as a linear function of the time of testing after the peak of instruction. For instance, if a group of individuals passed through a particular instructional sequence were tested at the end of instruction, and then at a later date were brought back to be tested again on the same material, a single linear transformation would be expected to equate each individual's scores on the two tests if the same unidimensional trait space was in operation at both times of testing. This would occur because the metric underlying the trait space would have retained its interval properties (Lord & Novick, 1968) with the passage of time and no further instruction would have occurred that might change the ordering of the individuals in terms of achievement level.

Further, if the same group were brought back for additional tests at later dates, a linear trend should be found in the comparison of any two testing periods, provided that the latent space did not change. If the latent space did vary, a linear relation between the two measures of achievement level would not be expected. Thus, if a linear relationship is not observed between scores obtained at and after the peak of instruction, the conclusion can be drawn that two different traits were being evaluated at the two testing times.

This study was concerned with determining whether the unidimensional space defining achievement at the peak of instruction was sufficient to describe achievement after the peak of instruction.

Method

Testing Procedure

Testing at the peak of instruction was a requirement for students enrolled in the same undergraduate survey course in biology as in Study 1. The peak-of-instruction test data were from the required first midquarter examinations administered in a 2-day period to all students enrolled in the course in the fall academic quarter of 1976 and in the winter quarter of 1977.

Testing after the peak of instruction was implemented by the Computerized Adaptive Testing Project using volunteers from the same biology classes. These volunteers were given extra points toward their final course grade for participating in the research and were told that their level of performance on the computer-administered biology test would have no effect on their final grades. This testing began on the day following the midquarter examination and continued for approximately 1 month.

Peak-of-instruction testing (the required first midquarter examination) consisted of the conventional paper-and-pencil administration of 55 multiple-choice questions concerning the first three content areas in the course--"Chemistry," "The Cell," and "Energy." (For a more complete discussion of the course content and testing procedure, see Bejar, Weiss, & Kingsbury, 1977.)

The after-peak-of-instruction test (the second, voluntary test) was a computer-administered stradaptive test (Weiss, 1973) consisting of a maximum of 50 individually selected items chosen from the same item pool that was used to construct the required midquarter, including the same three content areas. (For a complete description of this testing procedure, see Bejar, Weiss, & Gialluca, 1977). The data used for this study were from 253 students from the fall quarter testing for whom achievement estimates from both tests and the date of the later test were available.

Since the adaptive and conventional tests were selected from the same content area pools, these two tests should have measured the same underlying dimension if the passage of time did not affect the latent space; and although differences in the precision of measurement between the two testing procedures were present (Bejar, Weiss, & Gialluca, 1977), they should not have affected the outcome of the present study.

Scoring

Peak-of-instruction achievement level estimates were obtained by scoring the students' midquarter item response data with the scoring program LINDSCO (Bejar & Weiss, 1979), which is designed to score conventional tests using item characteristic curve models. After-peak-of-instruction achievement estimates were obtained by scoring the stradaptive response vectors with the program ADADSCO (Bejar & Weiss, 1979), which is designed to score adaptive tests with item characteristic curve models. Since the maximum-likelihood logistic scoring method used in both programs is the same, the achievement level estimates obtained from the two programs are directly comparable.

Analysis

To determine whether the same latent space was operative after instruction that was operative at the peak of instruction, individuals' achievement levels at the peak of instruction were regressed on their achievement estimates after the peak of instruction. Since the later testing occurred over a period of a month, it was possible to analyze the effect of the passage of time on the relationship between achievement estimates. Since after-peak-of-instruction testing occurred only on weekdays, the weekends served as natural break points to divide the total group of students into four subgroups, each of which was tested during a different week in the month following the peak of instruction. Table 4 shows the total number of students tested each quarter, as well as the number tested in each week following the first midquarter examination.

If the time of testing after peak of instruction affects the latent space underlying testee responses, this effect may be studied by examination of the regression lines of peak testing on later testing using data from each of the 4 weeks of testing. To the extent that these regressions are parallel (i.e., exhibit no interaction between achievement level at the peak of instruction and the time of after-peak-of-instruction testing) and exhibit stable linear

Table 4
Total Number of Students Tested Each Quarter and
Number Tested in Each Week Following Peak of Instruction
(First Midquarter Examination)

Group	Quarter: Fall	Quarter: Winter
Week 1	54	54
Week 2	83	90
Week 3	87	35
Week 4	29	6
Total	253	185

trends, it may be concluded that the latent space is stable and that the achievement metric was unchanged with the passage of time. As the time between testings lengthens, if increasing deviations from parallelism and/or linearity are observed, it may be concluded that the underlying trait space changed with the passage of time after instruction. Thus, both the parallelism and linearity of the regression of peak-of-instruction achievement level estimates on achievement level estimates obtained from after-peak-of-instruction testing were investigated.

Parallelism of regressions. For each of the 4 weeks of testing following the test administered at the peak of instruction, a separate regression line was obtained to predict individuals' later achievement levels from their peak-of-instruction achievement level estimates, using the subprogram REGRESSION contained in the Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1970). In addition, the overall regression line was obtained, including all individuals regardless of the date of the later testing.

To statistically examine the parallelism of the regression lines from the 4 weeks following the classroom examination, it was necessary to determine whether the individual lines fit the data any better than the single overall regression line. This analysis used a test statistic described by Neter and Wasserman (1974). The statistic which determines whether the full model (F; the four individual regression lines) substantially reduced the sum of squares due to error (SSE) in the restricted model (R; the single overall regression line) in this application is

$$W = \frac{SSE(R) - SSE(F)}{6} \Big/ \frac{SSE(F)}{N} \quad . \qquad [1]$$

where N equals the number of individuals tested.

W is distributed as an F (6, N-8) distribution, and a significant value implies that the full model of four individual regressions is significantly more precise than the single restricted model. If the value of W is not significant, it implies that predictions of the students' achievement levels at the peak of instruction are just as good if the week of the later testing is ignored.

If the value of the statistic in Equation 1 is statistically significant, the individual regression lines are different in some respect from the overall regression line; it is then appropriate to test directly whether the individual regression lines differ significantly in slope. This may again be done through the use of Equation 1. In this instance, the restricted model becomes a single regression equation predicting individuals' later achievement estimates from their achievement estimates at the peak of instruction and from the week of the later testing. The full model uses these two predictors and adds the interaction of the two predictors to the model. A significant value for the statistic indicates a significant interaction between the predictor variables, indicating that the individual regression lines are not parallel. A significant result from this analysis would indicate a change in the latent space.

These analyses were implemented for both the fall quarter testing group and the winter quarter testing group in order to examine the stability of the results across independent groups.

Polynomial trend analysis. For each week of testing following the peak of instruction, it was desired to determine whether a linear trend existed and was sufficient to describe the prediction of the after-peak-of-instruction achievement estimate from the achievement estimate obtained at the peak of instruction. This was operationalized by fitting a fourth-degree polynomial regression equation to the data for each week of testing and separately determining the significance of each of the terms in the equations. To the extent that these regression equations exhibited an increasing trend toward curvilinearity as the time between testings increased, it could be inferred that the latent space was changing with time, causing a disruption in the interval properties of the original metric. If no such trend was observed, it could be concluded that the latent space remained stable as a function of time.

Regression equations were obtained from the REGRESSION subprogram in the *Statistical Package for the Social Sciences* (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1970). Similar to the previous analysis, this analysis was implemented for both fall and winter quarters to permit replication of the results in independent groups.

Results

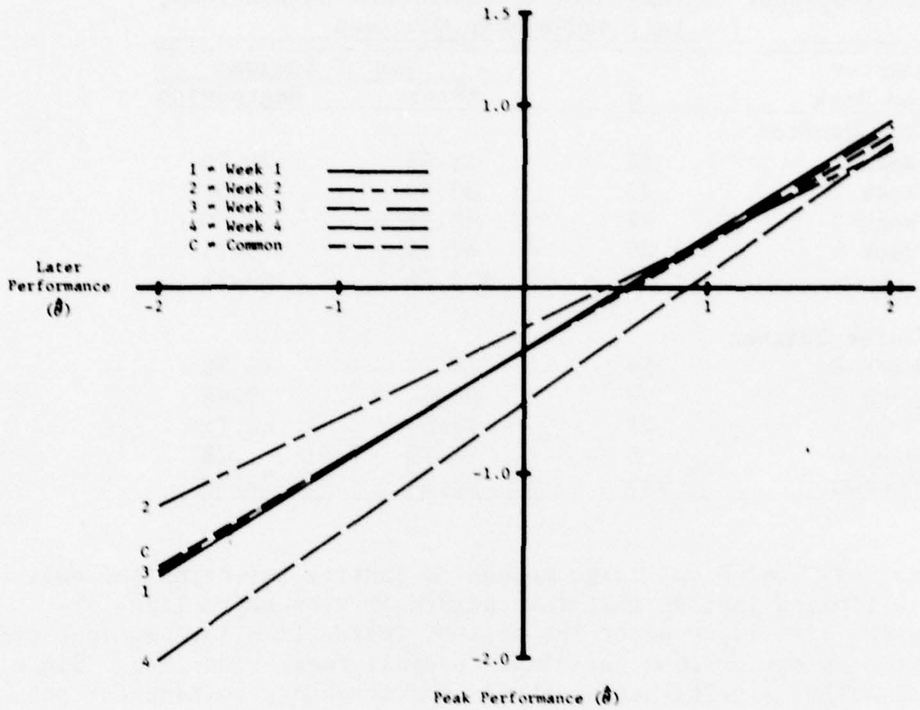
Parallelism of Regressions

Figure 5 shows individual regression lines obtained from each week of testing following the peak of instruction in the fall quarter (Figure 5a) and the winter quarter (Figure 5b), as well as the restricted regression line across weeks. Table 5 shows the sum of squares due to error and other descriptive statistics for each of the regression lines shown in Figure 5.

Using these sums of squares, the test for coincidence of regression in Equation 1 resulted in an F-value of 1.48 for the fall quarter data. This value, with 6 and 245 degrees of freedom, had a probability of occurrence by random fluctuation of $.10<p<.25$. For the winter quarter data, the observed F-value was .90, with 6 and 177 degrees of freedom. The probability of obtaining an F-value at least this extreme by random fluctuation was $p>.25$.

Figure 5
Regression Lines for Each Week Following Peak of Instruction
(First Midquarter Examination)

(a) Fall Quarter



(b) Winter Quarter

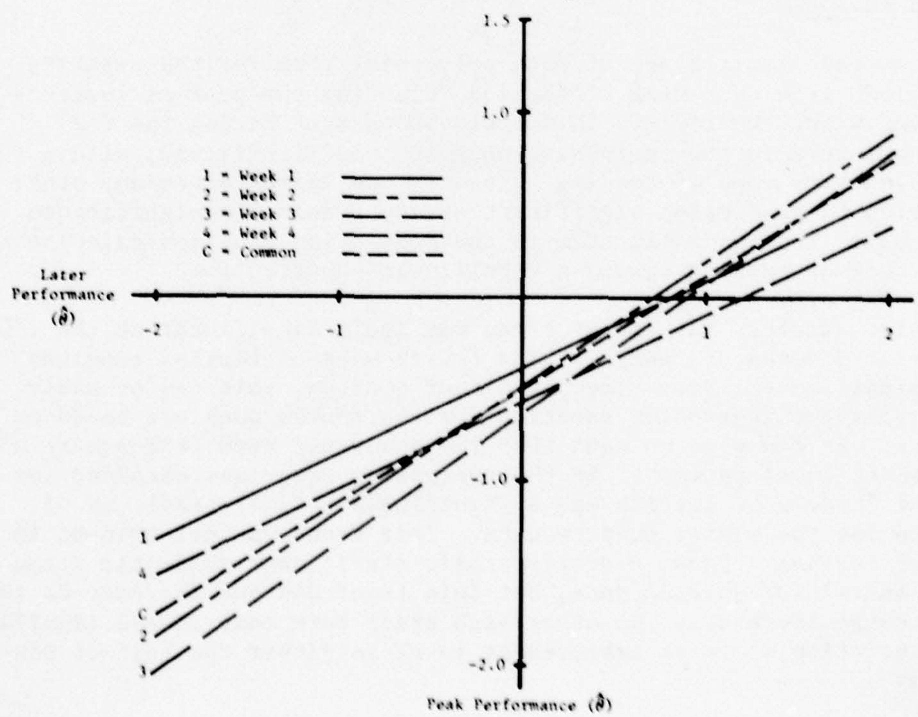


Table 5
Number of Subjects, Sum of Squares Due to Regression, and Sum of Squares Due to Error for Week-by-Week Regressions and for Common Regressions, for Fall and Winter Quarters

Quarter and Week	N	Sum of Squares: Error	Sum of Squares: Regression
Fall Quarter			
Week 1	54	21.88	23.76
Week 2	83	33.50	23.52
Week 3	87	41.32	36.70
Week 4	29	12.54	15.11
Common	253	113.24	103.26
Winter Quarter			
Week 1	54	20.77	13.32
Week 2	90	58.34	39.94
Week 3	35	12.83	20.75
Week 4	6	1.10	.78
Common	185	95.88	76.28

Neither obtained F-value was large enough to justify rejecting the null hypothesis. This finding implies that the individual regression lines obtained by taking the time lapse after the peak of instruction into account predicted later scores no better than the single overall regression line. Since the lines did not differ significantly, it was not necessary to test for parallelism among individual regression lines.

Polynomial Trend Analysis

Table 6 shows the significance of each polynomial term for the separate regression equations from each week of testing following the peak of instruction, for fall and winter quarters. This table shows that during the fall quarter, the linear term in the regression equation was significant, with a value of $p<.001$ in every week of testing. In only one instance was any other term's contribution to prediction significant at any reasonable significance level (i.e., $p<.05$): The quadratic term in the regression equation calculated from the fourth week of testing showed a significant contribution.

For the winter quarter the linear trend was again significant at the .001 level for the first 3 weeks of testing. The fourth week of testing resulted in the only nonsignificant linear trend in either quarter; this was probably due to the fact that the regression equation for the fourth week was based on only six students. It can also be seen that the quadratic term (the square of the peak achievement level estimate) in the regression equations obtained for each of the first 2 weeks of testing was a significant ($p<.05$) predictor of later performance for the winter quarter data. This trend was not evident in the third week of testing. Thus, a statistically significant quadratic trend was observed in the winter quarter data, but this trend did not increase as the time between testings increased. No other high order term contributed significantly to the prediction of later achievement level in either the fall or winter quarter data.

Table 6
Statistical Significance Level of Each Term of the Fourth Degree Polynomial Regression Equation Predicting Later Performance from Performance at the Peak of Instruction for Fall and Winter Quarters

Quarter and Week	N	Linear	Quadratic	Cubic	Quartic
Fall Quarter					
Week 1	54	$p<.001$	.358	.927	.693
Week 2	83	$p<.001$	.588	.326	.134
Week 3	87	$p<.001$	.510	.525	.518
Week 4	29	$p<.001$	.042	.748	.402
Winter Quarter					
Week 1	54	$p<.001$	.010	.671	.600
Week 2	90	$p<.001$	.018	.833	.589
Week 3	35	$p<.001$	.260	.541	.207
Week 4	6	.167	[a]	[a]	[a]

[a]Results not reported due to small N.

Discussion

Since the regression lines obtained in the different weeks of after-peak-of-instruction testing did not differ significantly from one another, the results of these analyses did not support the hypothesis that the achievement metric changed as a function of the time lapse between the peak of instruction and later achievement testing. The data indicate that the trait space was stable to the limit of the power of this analysis. It can also be seen from the data (see Figure 5) that student achievement measured after the peak of instruction was shifted in a linear manner from that measured at the peak of instruction; this is indicated by the nonzero intercepts and nonunit slopes of the overall regression lines from both quarters. The difference may be due to a lack of motivation or preparation for the after-peak-of-instruction tests, since that testing was voluntary and the scores on that test had no effect on students' course grades. In both quarters, however, there was no evidence of any change in the latent space with the passage of time.

The polynomial trend analysis indicated that in each week of testing following the peak of instruction, the maximum-likelihood estimates of achievement level at the peak of instruction were a significant linear predictor of later achievement levels. This finding was consistent across academic quarters. For the fall quarter the quadratic trend was significant only in the final week of testing. For the winter quarter the quadratic trend was a significant predictor in the first 2 weeks of testing, but not in the third week. These inconsistent findings imply that the quadratic trend observed may be a sample artifact.

The results from this analysis indicate that the only polynomial trend that acted as a consistent indicant of achievement was the linear term. It is, therefore, probable that the metric underlying individual testee response had not changed in any increasingly nonlinear manner, as would be expected if

the time between testings systematically affected the metric along which achievement was being measured, since no such systematic trend was noted.

Conclusions

The consistent findings of the analyses of the effect of time of after-peak-of-instruction testing on the measurement of achievement are as follows:

1. A single linear regression, using only prior performance as a predictor, was as efficient for the prediction of later performance as were four regression lines that took into account the time elapsed between peak of instruction and later testing.

2. The linear prediction trend was consistently significant in each week of testing following the peak of instruction.

3. No nonlinear prediction trend was consistently significant across all weeks of testing.

4. No significant increase in the significance of nonlinear prediction trends was observed with the increase of time elapsed between peak of instruction and later testing.

These findings lead to the conclusion that there was no evidence to support the hypothesis that the achievement variable changed as a function of the time lapse between peak of instruction and measurement of an individual's achievement level. The unidimensional ICC-based variable that had been used to measure achievement of individuals at the peak of instruction seemed to adequately describe the achievement of individuals as much as a month after the peak of instruction.

IMPLICATIONS AND LIMITATIONS

Implications

The results of these studies have implications for the measurement of achievement using both the pretest-test and test-posttest paradigms. The data suggest that there may be metric problems in the application of ICC item paramaters based on peak-of-instruction data to pretest data. Both the item difficulty and discrimination parameters estimated at the pretest (prior to instruction) differed substantially from those estimated at the peak of instruction. This result was reinforced by factor analyses of item sets obtained prior to instruction and at the peak of instruction; the variance accounted for by the first factor was considerably less at the pretest than it was at the peak of instruction. The implication of these results is that ICC-based pretest and peak-of-instruction achievement measurements may not be on the same dimension. Thus, the achievement variable measured at the pretest may be a different variable than that measured at the peak of instruction.

The importance of this finding, if it can be replicated in other data sets, is to call into question the utility of the pretest-test model for measuring gains in achievement. If the pretest achievement variable is, in fact, a different variable from that measured at the conclusion of instruction, it

is inappropriate to compute individual or group gain or change scores as indicants of growth in achievement. Such change scores would be completely useless in providing reliable estimates of growth in achievement levels because of the differences in the variables involved at the two points in time.

Contrary to the negative implications of these data for measuring achievement in a pretest-test paradigm, results of the second study support the use of an ICC-based test-posttest paradigm for the measurement of retention, at least within the 1-month time interval studied. Data from the second study showed that ICC-based achievement level estimates taken as much as a month after the peak of instruction were consistently linearly related to achievement level estimates taken at the peak of instruction. Thus, the data indicate that these posttest measurements were on the same ICC metric as the peak-of-instruction achievement level estimates. The data did show a level difference in the achievement estimates, but this might have resulted from design aspects of the study. Should future studies replicate this result (with or without the level difference), the data imply that gain (or loss) scores measuring retention after the peak of instruction may be meaningfully determined using ICC-based approaches.

The positive findings from the after-peak-of-instruction data also are in support of the potential of computerized adaptive testing for applications in the measurement of achievement. The data indicated linear relationships among ICC-based achievement level estimates obtained up to 4 weeks after instruction. Thus, even with limited availability of testing terminals, which might be characteristic of adaptive testing in certain instructional environments, it may be possible to obtain equivalent achievement estimates for students tested as long as a few weeks after the material was covered in a course. This should minimize the cost of an adaptive testing system and make its use economically feasible for classrooms of all sizes. Further research will be necessary, of course, to determine whether the observed mean differences in student performance after the peak of instruction were due to the motivational factors characteristic of voluntary participation.

Limitations

Both of the studies reported above were done within the context of a single undergraduate survey biology course. This limits the generalizability of the studies in several ways. For example, the weak factor structure noted in students' responses prior to instruction may be due to the fact that this was an introductory course. It is very possible that a more advanced course might show a strong prior-to-peak-of-instruction factor structure which is similar to the peak-of-instruction factor structure. In addition, different items drawn from the same content pool were used in the prior-to-instruction and peak-of-instruction factor comparisons. Future studies should compare the factor structures of the same items prior to instruction and at the peak of instruction.

Further limitations in the constancy of the testing procedures might have added some biases to the conclusions drawn. The pretest measure was required of all students attending the first lecture of the class, but it did not affect students' grades in the course. The first midquarter examination (peak-of-instruction test) was required of all students in the course and did have a bearing on the students' grade in the course. The after-peak-of-instruction measure was a voluntary test which allowed students to add extra credit points

to their course grade. In addition, this test was a computer-administered stradaptive test, whereas the first two tests were administered in conventional paper-and-pencil format. These differences, both methodological and motivational, may have added some unknown amount of bias to the results of the studies. Thus, replication of these studies in other instructional environments, and with revisions in the research design, is appropriate.

REFERENCES

Bejar, I. I., & Weiss, D. J. A construct validation of adaptive achievement testing (Research Report 78-4). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, November 1978.

Bejar, I. I., & Weiss, D. J. Computer programs for scoring test data with item characteristic curve models (Research Report 79-1). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, February 1979. (NTIS No. AD A067928).

Bejar, I. I., Weiss, D. J., & Gialluca, K. A. An information comparison of conventional and adaptive tests in the measurement of classroom achievement (Research Report 77-7). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, October 1977. (NTIS No. AD A047495).

Bejar, I. I., Weiss, D. J., & Kingsbury, G. G. Calibration of an item pool for the adaptive measurement of achievement (Research Report 77-5). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, September 1977. (NTIS No. AD A044828).

Brown, J. M., & Weiss, D. J. An adaptive testing strategy for achievement test batteries (Research Report 77-6). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, September 1977. (NTIS No. AD A046062).

Cronbach, L. J., & Furby, L. How we should measure "change"--or should we? Psychological Bulletin, 1970, 74, 68-80.

Harris, C. W. (Ed.) Problems in measuring change. Madison: University of Wisconsin Press, 1963.

Kingsbury, G. G., & Weiss, D. J. Relationship among achievement level estimates from three item characteristic curve scoring methods (Research Report 79-3). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, April 1979.

Lord, F. M., & Novick, M. R. Statistical theories of mental test scores. Reading, MA: Addison-Wesley, 1968.

McBride, J. R., & Weiss, D. J. Some properties of a Bayesian adaptive ability testing strategy (Research Report 76-1). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, March 1976. (NTIS No. AD A022964).

Neter, J., & Wasserman, W. Applied linear statistical models. Homewood, IL: Richard D. Irwin, 1974.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. Statistical package for the social sciences. New York: McGraw-Hill, 1970.

Urry, U. W. A five-year quest: Is computerized adaptive testing feasible? In C. L. Clark (Ed.), Proceedings of the first conference on computerized adaptive testing (U.S. Civil Service Commission, Personnel Research and Development Center, PS-75-6). Washington, DC: U.S. Government Printing Office, 1976. (Superintendent of Documents Stock No. 006-00940-9).

Vale, C. D. Problem: Strategies of branching through an item pool. In D. J. Weiss (Ed.), Computerized adaptive trait measurement: Problems and prospects (Research Report 75-5). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, November 1975. (NTIS No. AD A018675).

Vale, C. D., & Weiss, D. J. A simulation study of stradaptive ability testing (Research Report 75-6). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, December 1975. (NTIS No. AD A020961).

Weiss, D. J. The stratified adaptive computerized ability test (Research Report 73-3). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, September 1973. (NTIS No. AD 768376).

APPENDIX: SUPPLEMENTARY TABLES

Table A
Parameter Estimates for Items Used in
Study of Peak/Peak Parameter Correspondence

Item Number	First Administration Test*	Parameter *a*	*b*	*c*	Second Administration Test*	Parameter *a*	*b*	*c*
3002	WF	.82	.13	.14	SF	.87	.12	.27
3034	W1	1.01	.37	.28	S1	.85	-.29	.13
3038	W1	1.58	-.56	.28	S1	1.20	-1.06	.16
3201	W1	1.07	-1.34	.23	S1	.85	-1.74	.18
3206	W1	.74	1.51	.21	S1	.75	1.57	.32
3216	W1	1.27	-.62	.18	S1	1.17	-.60	.15
3218	W1	.82	.58	.12	S1	.80	.34	.14
3237	WF	1.54	-.37	.18	SF	1.58	-.11	.43
3241	W1	1.12	2.48	.24	S1	.91	2.09	.17
3414	W1	.88	2.29	.32	S1	1.40	1.96	.30
3651	W2	.81	2.27	.44	S2	.95	2.31	.52
3812	W2	.74	-.66	.11	S2	.82	-.63	.13
3909	W2	1.34	.77	.38	S2	.90	1.12	.36
4006	WF	.84	-.59	.16	SF	1.05	-.19	.27
4036	WF	1.24	-.61	.23	SF	.95	-1.30	.17
4044	WF	.80	-.12	.38	SF	.80	-.60	.13
4229	WF	1.36	-.45	.38	SF	1.64	-.92	.17
4238	WF	.83	1.54	.42	SF	.83	1.47	.43

*W=Winter Quarter 1976; S=Spring Quarter 1976; 1=First Midquarter Examination; 2=Second Midquarter Examination; F=Final Examination

Table B

Interitem Correlations among 21 Items Selected from Pretest (Lower Triangle) and First Midquarter Examination (Upper Triangle)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	--	23	32	21	28	03	18	06	20	20	39	39	18	40	34	27	08	26	32	16	22
2	11	--	25	14	13	16	16	12	15	20	16	09	15	26	22	27	08	16	25	11	03
3	-11	02	--	15	28	22	22	16	23	16	22	24	06	15	29	31	12	12	29	12	11
4	04	-02	13	--	15	14	13	06	08	19	14	23	-02	20	14	24	19	12	18	15	11
5	11	00	01	05	--	55	25	34	36	37	39	21	21	26	38	25	23	12	39	13	21
6	01	-03	05	07	05	--	18	21	28	21	23	19	19	21	25	27	20	22	39	-04	10
7	02	06	03	10	00	10	--	11	23	00	19	15	15	13	13	19	10	06	29	13	12
8	07	-07	09	12	05	12	17	--	33	22	26	22	15	16	20	28	16	11	30	07	15
9	00	01	-05	-04	12	01	-02	04	--	27	30	19	24	08	21	28	19	19	31	14	17
10	07	02	03	03	18	05	04	01	08	--	26	15	27	35	23	34	14	15	38	17	10
11	19	01	01	06	12	-03	08	23	08	09	--	26	19	33	35	33	14	27	40	09	20
12	11	-03	03	12	05	00	08	11	-02	04	16	--	14	13	18	28	23	22	22	12	10
13	17	06	15	57	14	13	20	24	-01	04	22	14	--	26	24	15	-01	20	30	02	06
14	-04	03	10	05	-02	04	01	02	-05	06	03	-03	11	--	22	16	06	12	42	-01	11
15	24	-02	03	19	13	01	15	23	02	-02	22	09	24	09	--	36	13	22	38	-05	22
16	06	00	10	08	04	00	11	12	04	12	08	14	11	03	03	--	19	28	45	19	21
17	-01	00	01	04	-10	14	01	04	-10	-01	03	02	06	-01	-03	-13	--	16	19	19	16
18	20	01	08	21	06	00	19	14	-01	07	06	17	21	05	11	17	08	--	29	12	05
19	03	05	01	04	07	08	18	05	-09	-06	05	03	00	01	04	10	02	09	--	03	24
20	03	07	04	04	09	-02	01	10	00	07	05	10	11	16	07	-01	06	08	-02	--	15
21	19	03	02	09	07	06	16	03	06	-04	10	07	15	-03	11	-05	06	19	-01	02	--

Note. Item numbers are arbitrary; different items were selected from the same content areas at both Pretest and First Midquarter.

Table C
Five-Factor Solutions Prior to Instruction (Pretest) and at Peak of Instruction (First Midquarter Examination)

Variable	Factor 1	2	3	4	5
Prior to Instruction					
1	.32	.38	-.26	.00	-.18
2	.04	.05	-.06	.01	-.13
9	.16	-.18	.18	-.01	.03
11	.54	-.40	-.03	-.11	-.12
15	.21	.19	-.01	-.22	-.01
17	.14	-.09	-.00	.13	.09
18	.32	.03	.10	.29	-.01
19	.41	.08	.14	.11	.41
22	.02	.17	.01	-.19	.04
23	.12	.12	.12	-.18	-.05
24	.36	.26	-.04	-.09	.19
25	.27	.11	.08	.04	-.03
27	.75	-.32	-.11	-.18	-.03
29	.11	-.10	.07	-.05	.06
30	.41	.15	-.12	-.03	.15
32	.25	.15	.51	-.04	-.17
33	.06	-.14	-.15	.24	.08
36	.41	.08	.07	.20	-.22
37	.12	.05	.12	.24	-.04
38	.16	.01	-.01	-.06	.07
39	.25	.11	-.24	.15	-.11
Peak of Instruction					
1	.59	.68	-.08	-.33	.01
4	.36	.11	-.04	.16	.00
7	.44	.07	.10	-.03	-.15
10	.31	.12	.15	.03	.07
13	.70	-.46	-.04	-.48	.15
14	.49	-.38	-.06	-.03	-.14
16	.33	-.01	.09	-.07	-.11
19	.41	-.20	.07	.07	-.01
22	.49	-.16	.13	.03	-.02
25	.51	-.07	-.10	.23	.40
28	.58	.07	-.05	-.04	-.04
31	.43	.15	.19	-.05	-.10
34	.36	-.04	-.22	.12	.04
37	.48	.16	-.39	.05	.20
40	.54	.01	-.13	-.00	-.19
41	.60	.04	.20	.28	-.11
43	.31	-.08	.30	.00	.02
46	.38	.10	.06	.14	-.11
49	.69	-.05	-.16	.19	-.08
52	.20	.12	.46	.01	.32
55	.31	.03	.12	-.07	-.01

DISTRIBUTION LIST

Navy

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code N-71
NAVTRAEQUIPCEN
Orlando, FL 32813

1 MR. MAURICE CALLAHAN
Pers 23a
Bureau of Naval Personnel
Washington, DC 20370

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Dr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 CAPT. D.M. GRAGG, MC, USN
HEAD, SECTION ON MEDICAL EDUCATION
UNIFORMED SERVICES UNIV. OF THE
HEALTH SCIENCES
6917 ARLINGTON ROAD
BETHESDA, MD 20014

1 CDR Robert S. Kennedy
Naval Aerospace Medical and
Research Lab
Box 29407
New Orleans, LA 70189

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

1 CHAIRMAN, LEADERSHIP & LAW DEPT.
DIV. OF PROFESSIONAL DEVELOPMMENT
U.S. NAVAL ACADEMYY
ANNAPOLIS, MD 21402

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 CAPT Richard L. Martin
USS Francis Marion (LPA-Z49)
FPO New York, NY 09501

1 Dr. James McBride
Code 301
Navy Personnel R&D Center
San Diego, CA 92152

2 Dr. James McGrath
Navy Personnel R&D Center
Code 306
San Diego, CA 92152

1 DR. WILLIAM MONTAGUE
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Commanding Officer
Naval Health Research
Center
Attn: Library
San Diego, CA 92152

1 Naval Medical R&D Command
Code 44
National Naval Medical Center
Bethesda, MD 20014

1 Library
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

1 OFFICE OF CIVILIAN PERSONNEL
(CODE 26)
DEPT. OF THE NAVY
WASHINGTON, DC 20390

1 JOHN OLSEN
CHIEF OF NAVAL EDUCATION &
TRAINING SUPPORT
PENSACOLA, FL 32509

1 Psychologist
ONR Branch Office
495 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Office of Naval Research
Code 200
Arlington, VA 22217

1 Code 436
Office of Naval Research
Arlington, VA 22217

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
OFFICE OF NAVAL RESEARCH BRANCH
223 OLD MARYLEBONE ROAD
LONDON, NW, 15TH ENGLAND

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Scientific Director
Office of Naval Research
Scientific Liaison Group/Tokyo
American Embassy
APO San Francisco, CA 96503

1 Office of the Chief of Naval Operations
Research, Development, and Studies Branc
(OP-102)
Washington, DC 20350

1 Scientific Advisor to the Chief of
Naval Personnel (Pers-Or)
Naval Bureau of Personnel
Room 4410, Arlington Annex
Washington, DC 20370

1 LT Frank C. Petho, MSC, USNR (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

1 DR. RICHARD A. POLLAK
ACADEMIC COMPUTING CENTER
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Roger W. Remington, Ph.D
Code L52
NAMRL
Pensacola, FL 32508

1 Dr. Bernard Rimland
Navy Personnel R&D Center
San Diego, CA 92152

1 Mr. Arnold Rubenstein
Naval Personnel Support Technology
Naval Material Command (08T244)
Room 1044, Crystal Plaza #5
2221 Jefferson Davis Highway
Arlington, VA 20360

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

1 A. A. SJOHOLM
TECH. SUPPORT, CODE 201
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

1 Mr. Robert Smith
Office of Chief of Naval Operations
OP-987E
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 CDR Charles J. Theisen, JR. MSC, USN
Head Human Factors Engineering Div.
Naval Air Development Center
Warminster, PA 18974

1 W. Gary Thomson
Naval Ocean Systems Center
Code 7132
San Diego, CA 92152

1 Dr. Ronald Weitzman
Department of Administrative Sciences
U. S. Naval Postgraduate School
Monterey, CA 93940

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 HQ USAREUE & 7th Army
ODCSOPS
USAAREUE Director of GED
APO New York 09403

1 LCOL Gary Bloedorn
Training Effectiveness Analysis Division
US Army TRADOC Systems Analysis Activity
White Sands Missile Range, NM 88002

1 DR. RALPH DUSEK
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Ed Johnson
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

1 Dr. Michael Kaplan
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Milton S. Katz
Individual Training & Skill
Evaluation Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Beatrice J. Farr
Army Research Institute (PERI-OK)
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Milt Maier
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Harold F. O'Neil, Jr.
ATTN: PERI-OK
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Robert Ross
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Director, Training Development
U.S. Army Administration Center
ATTN: Dr. Sherrill
Ft. Benjamin Harrison, IN 46218

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab
AFHRL/PED
Brooks AFB, TX 78235

1 Air University Library
AUL/LSE 76/443
Maxwell AFB, AL 36112

1 Dr. Philip De Leo
AFHRL/TT
Lowry AFB, CO 80230

1 DR. G. A. ECKSTRAND
AFHRL/AS
WRIGHT-PATTERSON AFB, OH 45433

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 CDR. MERCER
CNET LIAISON OFFICER
AFHRL/FLYING TRAINING DIV.
WILLIAMS AFB, AZ 85224

1 Dr. Ross L. Morgan (AFHRL/ASR)
Wright -Patterson AFB
Ohio 45433

1 Dr. Roger Pennell
AFHRL/TT
Lowry AFB, CO 80230

1 Personnel Analysis Division
HQ USAF/DPXXA
Washington, DC 20330

1 Research Branch
AFMPC/DPMYP
Randolph AFB, TX 78148

1 Dr. Malcolm Ree
AFHRL/PED
Brooks AFB, TX 78235

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

CoastGuard

1 Mr. Richard Lanterman
PSYCHOLOGICAL RESEARCH (G-P-1/62)
U.S. COAST GUARD HQ
WASHINGTON, DC 20590

1 Dr. Thomas Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Other DoD

12 Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314
Attn: TC

1 Dr. Dexter Fletcher
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD.
ARLINGTON, VA 22209

1 Dr. William Graham
Testing Directorate
MEPCOM
Ft. Sheridan, IL 60037

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 MAJOR Wayne Sellman, USAF
Office of the Assistant Secretary
of Defense (MRA&L)
3B930 The Pentagon
Washington, DC 20301

Civil Govt

1 Dr. Susan Chipman
Basic Skills Program
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. William Gorham, Director
Personnel R&D Center
Office of Personnel Managment
1900 E Street NW
Washington, DC 20415

1 Dr. Joseph I. Lipson
Division of Science Education
Room W-638
National Science Foundation
Washington, DC 20550

1 Dr. John Mays
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Arthur Melmed
National Intitute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Dr. Lalitha P. Sanathanan
Environmental Impact Studies Division
Argonne National Laboratory
9700 S. Cass Avenue
Argonne, IL 60439

1 Dr. Jeffrey Schiller
National Institute of Education
1200 19th St. NW
Washington, DC 20208

1 Dr. Thomas G. Sticht
Basic Skills Program
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Managment
1900 E Street NW
Washington, DC 20415

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Dr. Erling B. Anderson
University of Copenhagen
Studiestraedt
Copenhagen
DENMARK

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. Warner Birice
Streitkraefteamt
Rosenberg 5300
Bonn, West Germany D-5300

1 Dr. R. Darrel Bock
Department of Education
University of Chicago
Chicago, IL 60637

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. David G. Bowers
Institute for Social Research
University of Michigan
Ann Arbor, MI 48106

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52240

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. John Chiorini
Litton-Mellonics
Box 1286
Springfield, VA 22151

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. William Coffman
Iowa Testing Programs
University of Iowa
Iowa City, IA 52242

1 Dr. Allan M. Collins
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, Ma 02138

1 Dr. Meredith Crawford
Department of Engineering Administration
George Washington University
Suite 805
2101 L Street N. W.
Washington, DC 20037

1 Dr. Hans Cronbag
Education Research Center
University of Leyden
Boerhaavelaan 2
Leyden
The NETHERLANDS

1 MAJOR I. N. EVONIC
CANADIAN FORCES PERS. APPLIED RESEARCH
1107 AVENUE ROAD
TORONTO, ONTARIO, CANADA

1 Dr. Leonard Feldt
Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Gerhardt Fischer
Liebigasse 5
Vienna 1010
Austria

1 Dr. Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

1 Dr. Edwin A. Fleishman
Advanced Research Resources Organ.
Suite 900
4330 East West Highway
Washington, DC 20014

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Ross Greene
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

1 Dr. Alan Gross
Center for Advanced Study in Education
City University of New York
New York, NY 10036

1 Dr. Ron Hambleton
School of Education
University of Massechusetts
Amherst, MA 01002

1 Dr. Chester Harris
School of Education
University of California
Santa Barbara, CA 93106

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

1 Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Huynh Huynh
Department of Education
University of South Carolina
Columbia, SC 29208

1 Dr. Carl J. Jensema
Gallaudet College
Kendall Green
Washington, DC 20002

1 Dr. Arnold F. Kanarick
Honeywell, Inc.
2600 Ridgeway Pkwy
Minneapolis, MN 55413

1 Dr. John A. Keats
University of Newcastle
Newcastle, New South Wales
AUSTRALIA

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 LCOL. C.R.J. LAFLEUR
PERSONNEL APPLIED RESEARCH
NATIONAL DEFENSE HQS
101 COLONEL BY DRIVE
OTTAWA, CANADA K1A 0K2

1 Dr. Michael Levine
Department of Educational Psychology
University of Illinois
Champaign, IL 61820

1 Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Dr. Frederick M. Lord
Educational Testing Service
Princeton, NJ 08540

1 Dr. Robert R. Mackie
Human Factors Research, Inc.
6780 Cortona Drive
Santa Barbara Research Pk.
Goleta, CA 93017

1 Dr. Gary Marco
Educational Testing Service
Princeton, NJ 08450

1 Dr. Scott Maxwell
Department of Psychology
University of Houston
Houston, TX 77025

1 Dr. Sam Mayo
Loyola University of Chicago
Chicago, IL 60601

1 Dr. Allen Munro
Univ. of So. California
Behavioral Technology Labs
3717 South Hope Street
Los Angeles, CA 90007

1 Dr. Melvin R. Novick
Iowa Testing Programs
University of Iowa
Iowa City, IA 52242

1 Dr. Jesse Orlansky
Institute for Defense Analysis
400 Army Navy Drive
Arlington, VA 22202

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. STEVEN M. PINE
4950 Douglas Avenue
Golden Valley, MN 55416

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

1 MIN. RET. M. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 161
53 BONN 1, GERMANY

1 Dr. Peter B. Read
Social Science Research Council
605 Third Avenue
New York, NY 10016

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
12 Hill Hall
Columbia, MO 65201

1 Dr. Fred Reif
SESAME
c/o Physics Department
University of California
Berkely, CA 94720

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 Dr. Donald Rubin
Educational Testing Service
Princeton, NJ 08450

1 Dr. Larry Rudner
Gallaudet College
Kendall Green
Washington, DC 20002

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Dr. Kazao Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 982
JAPAN

1 Dr. Edwin Shirkey
Department of Psychology
Florida Technological University
Orlando, FL 32816

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 DR. ALBERT STEVENS
BOLT BERANEK & NEWMAN, INC.
50 MOULTON STREET
CAMBRIDGE, MA 02138

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN
THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Brad Sympson
Office of Data Analysis Research
Educational Testing Service
Princeton, NJ 08541

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 Dr. Maurice Tatsuoka
Department of Educational Psychology
University of Illinois
Champaign, IL 61801

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Robert Tsutakawa
Dept. of Statistics
University of MIssouri
Columbia, MO 65201

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

1 Dr. Howard Wainer
Bureau of Social SCience Research
1990 M Street, N. W.
Washington, DC 20036

1 DR. THOMAS WALLSTEN
PSYCHOMETRIC LABORATORY
DAVIE HALL 013A
UNIVERSITY OF NORTH CAROL
CHAPEL HILL, NC 27514

1 Dr. John Wannous
Department of Management
Michigan University
East Lansing, MI 48824

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

1 Dr. Wolfgang Wildgrube
Streitkraefteamt
Rosenberg 5300
Bonn, West Germany D-5300

1 Dr. Robert Woud
School Examination Department
University of London
66-72 Gower Street
London WCIE 6EE
ENGLAND

1 Dr. Karl Zinn
Center for research on Learning and Teaching
University of Michigan
Ann Arbor, MI 48104

Previous Publications

Proceedings of the 1977 Computerized Adaptive Testing Conference. July 1978.

Research Reports

79-3. Relationships among Achievement Level Estimates from Three Item Characteristic Curve Scoring Methods. April 1979.
Final Report: Bias-Free Computerized Testing. March 1979. (NTIS No. AD A068176)
79-2. Effects of Computerized Adaptive Testing on Black and White Students. March 1979. (NTIS No. AD A067928)
79-1. Computer Programs for Scoring Test Data with Item Characteristic Curve Models. February 1979. (NTIS No. AD A067752)
78-5. An Item Bias Investigation of a Standardized Aptitude Test. December 1978. (NTIS No. AD A064352)
78-4. A Construct Validation of Adaptive Achievement Testing. November 1978.
78-3. A Comparison of Levels and Dimensions of Performance in Black and White Groups on Tests of Vocabulary, Mathematics, and Spatial Ability. October 1978. (NTIS No. AD A062797)
78-2. The Effects of Knowledge of Results and Test Difficulty on Ability Test Performance and Psychological Reactions to Testing. September 1978.
78-1. A Comparison of the Fairness of Adaptive and Conventional Testing Strategies. August 1978. (NTIS No. AD A059436)
77-7. An Information Comparison of Conventional and Adaptive Tests in the Measurement of Classroom Achievement. October 1977. (NTIS No. AD A047495)
77-6. An Adaptive Testing Strategy for Achievement Test Batteries. October 1977. (NTIS No. AD A046062)
77-5. Calibration of an Item Pool for the Adaptive Measurement of Achievement. September 1977. (NTIS No. AD A044828)
77-4. A Rapid Item-Search Procedure for Bayesian Adaptive Testing. May 1977. (NTIS No. AD A041090)
77-3. Accuracy of Perceived Test-Item Difficulties. May 1977. (NTIS No. AD A041084)
77-2. A Comparison of Information Functions of Multiple-Choice and Free-Response Vocabulary Items. April 1977.
77-1. Applications of Computerized Adaptive Testing. March 1977. (NTIS No. AD A038114)
Final Report: Computerized Ability Testing, 1972-1975. April 1976. (NTIS No. AD A024516)
76-5. Effects of Item Characteristics on Test Fairness. December 1976. (NTIS No. AD A035393)
76-4. Psychological Effects of Immediate Knowledge of Results and Adaptive Ability Testing. June 1976. (NTIS No. AD A027170)
76-3. Effects of Immediate Knowledge of Results and Adaptive Testing on Ability Test Performance. June 1976. (NTIS No. AD A028147)
76-2. Effects of Time Limits on Test-Taking Behavior. April 1976. (NTIS No. AD A024422)
76-1. Some Properties of a Bayesian Adaptive Ability Testing Strategy. March 1976. (NTIS No. AD A022964)
75-6. A Simulation Study of Stradaptive Ability Testing. December 1975. (NTIS No. AD A020961)
75-5. Computerized Adaptive Trait Measurement: Problems and Prospects. November 1975. (NTIS No. AD A018675)
75-4. A Study of Computer-Administered Stradaptive Ability Testing. October 1975. (NTIS No. AD A018758)
75-3. Empirical and Simulation Studies of Flexilevel Ability Testing. July 1975. (NTIS No. AD A013185)
75-2. TETREST: A FORTRAN IV Program for Calculating Tetrachoric Correlations. March 1975. (NTIS No. AD A007572)
75-1. An Empirical Comparison of Two-Stage and Pyramidal Adaptive Ability Testing. February 1975. (NTIS No. AD A006733).
74-5. Strategies of Adaptive Ability Measurement. December 1974. (NTIS No. AD A004270)
74-4. Simulation Studies of Two-Stage Ability Testing. October 1974. (NTIS No. AD A001230)
74-3. An Empirical Investigation of Computer-Administered Pyramidal Ability Testing. July 1974. (NTIS No. AD 783553)
74-2. A Word Knowledge Item Pool for Adaptive Ability Measurement. June 1974. (NTIS No. AD 781894)
74-1. A Computer Software System for Adaptive Ability Measurement. January 1974. (NTIS No. AD 773961)
73-3. The Stratified Adaptive Computerized Ability Test. September 1973. (NTIS No. AD 768376)
73-2. Comparison of Four Empirical Item Scoring Procedures. August 1973.
73-1. Ability Measurement: Conventional or Adaptive? February 1973. (NTIS No. AD 757788)

AD Numbers are those assigned by the Defense Documentation Center, for retrieval through the National Technical Information Service.

Copies of these reports are available, while supplies last, from:
Psychometric Methods Program, Department of Psychology
N660 Elliott Hall, University of Minnesota
75 East River Road, Minneapolis, Minnesota 55455